A Loudspeaker Response Interpolation Model based on One-twelfth Octave Interval Frequency Measurements

Anthony Zaknich

School of Engineering, Murdoch University, Rockingham, 6168, Australia

ABSTRACT

A practical loudspeaker frequency response interpolation model is developed using a modification of the Tuneable Approximate Piecewise Linear Regression (TAPLR) model that can provide a complete magnitude and phase response over the full frequency range of the loudspeaker. This is achieved by first taking standard one-twelfth octave frequency interval acoustic intensity measurements at a one meter distance in front of the loudspeaker. These measurements are inserted directly into the formulation, which then requires only minimal tuning to achieve a magnitude response model to better than +/- 1 dB error as compared with the magnitude of the Fourier transform of the impulse response for typical hi-fi loudspeakers. The Hilbert transform can then be used to compute the corresponding phase response directly from the resulting magnitude response. Even though it is initially based on consecutive piecewise linear sections this new model provides a continuous smooth interpolation between the measured values that is much more satisfactory than normal piecewise linear segment interpolation and much simpler to do than polynomial interpolation. It only requires the tuning of a single parameter to control the degree of smoothness from a stair step response at one extreme to a straight mean horizontal line at the other. It is easy to find the best tuning parameter value in between these two extremes by either trial and error or by the minimisation of a mean squared interpolation error.

INTRODUCTION

For real loudspeakers systems, a sufficiently accurate frequency response model is difficult, if not impossible, to fully compute from theory and therefore needs to be determined from measured data. Once appropriate measurements of frequency and phase response have been made they then need to be encapsulated in a model of some sort; a model that will provide a complete and adequate response over the full frequency range of the loudspeaker. This paper shows that it is sufficient to take magnitude response measurements at the standard one-twelfth octave frequency intervals from a one meter distance in front of the loudspeaker to provide such a model that is accurate to better than +/- 1 dB for typical hi-fi loudspeakers. It is possible to achieve this by using a simple interpolation structure based on a slight modification of the Tuneable Approximate Piecewise Linear Regression (TAPLR) model (Zaknich and Attikiouzel 2000) in conjunction with the Hilbert transform (Poularikas 1996) to compute the relevant phase response from the subsequent magnitude response model.

The TAPLR model was derived from the Modified Probabilistic Neural Network (MPNN) (Zaknich 1998), which uses Radial Basis Functions (RBFs) to interpolate over a data space for the solution of regression problems. It is a generalization of Specht's General Regression Neural Network (GRNN) (Specht 1991) and both are similar to the method of (Moody and Darken 1989).

This paper provides a review of the general TAPLR model and its previous application to loudspeaker response modelling (Zaknich 2004). It completes and extends this previous work by applying and testing the model modification suggested in (Zaknich 2004) for the development of a new and better loudspeaker frequency response model. This new model is fully developed and tested using the same application example to demonstrate its improvements and advantages over the previous model.

REVIEW OF THE GENERAL TAPLR MODEL

The MPNN model was designed for general nonlinear regression problems given a set of input/output data pairs. However, it is not difficult to generalize the basic structure of the MPNN to combine and integrate any set of complementary local linear models, rather than single point data pairs to cover an input data space. The TAPLR model is defined by equation (1) and shown in Figure1.

$$\hat{y}(\mathbf{x}) = \frac{\sum_{i=1}^{M} Z_i l_i(\mathbf{x}) f(\|\mathbf{x}-\mathbf{c}_i\|,\sigma)}{\sum_{i=1}^{M} Z_i f(\|\mathbf{x}-\mathbf{c}_i\|,\sigma)} \tag{1}$$

where:

- $f(\|\mathbf{x}-\mathbf{c}_i\|,\sigma)$ is a suitable RBF, typically Gaussian.
- $\mathbf{x}$ is an arbitrary input space vector.
- $\mathbf{c}_i$ is the centre-vector for RBF i in the input space.
- M is the number of unique RBF centre-vectors $\mathbf{c}_i$.
- σ is the single smoothing parameter for model tuning.
- $l_i(\mathbf{x})$ is local linear model output associated with RBF i.
- Z_i is the fixed weight associated with RBF i.

If all the local linear models $l_i(\mathbf{x})$ are equally likely, or the a priori likelihood is unknown, and the centres of the RBFs are uniformly distributed in the data space then all the $Z_i = 1$. Otherwise, the relative values of Z_i may be set to represent the relative a priori likelihood or required relative weighting of each $l_i(\mathbf{x})$ as appropriate. The functional values given by each RBF weighted by Z_i, i.e. $Z_i f(\|\mathbf{x}-\mathbf{c}_i\|,\sigma)$, are used as a measure of closeness of $\mathbf{x}$ to each of the centre-vectors $\mathbf{c}_i$ (associated with each of the $l_i(\mathbf{x})$ models) and thus provide the required relative $l_i(\mathbf{x})$ model weightings. The TAPLR model is essentially a mixture model and as such further piecewise linear models $l_i(\mathbf{x})$ can be added to the structure to accommodate required design changes or specifications. The degree of local model coupling or decoupling can be simply controlled by the adjustment of the common RBF bandwidth parameter σ.

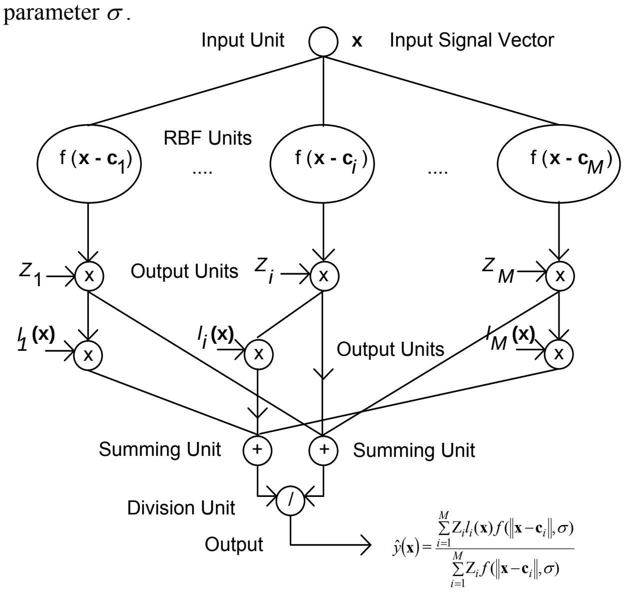

Figure 1. The TAPLR Model Architecture

A Gaussian radial (spherical) basis function, defined by equation (2), is often used for $f(\|\mathbf{x}-\mathbf{c}_i\|,\sigma)$. Adjustment of σ controls the degree of weighting of each linear local model associated with each centre-vector. Input vectors **x** closest to a centre-vector activate the associated linear model more than for those further away. For very small σ the linear model associated with the centre-vector closest to the current input point dominates, resulting in a linear response in the local space of that centre-vector. For very large σ the network output approaches a fixed weighted average of all the linear models. Somewhere in between a suitable overall model results, which provides approximately linear operation close to each centre-vector while deviating from linearity close to centre-vector region boundaries. At the boundaries between centre-vectors a smooth and continuous merging of neighbouring linear models occurs. The location of each centre-vector is usually chosen manually depending on the requirements of the problem; however it is a subject of ongoing research to automate this selection.

$$f(\|\mathbf{x}-\mathbf{c}_i\|,\sigma) = \exp\frac{-(\mathbf{x}-\mathbf{c}_i)^T(\mathbf{x}-\mathbf{c}_i)}{2\sigma^2} \tag{2}$$

Model tuning is done by finding the single best tuning parameter σ that results in the minimum Mean Squared Error (MSE) of the output $\hat{y}(\mathbf{x}_k)$ minus the corresponding desired output y_k for a representative testing set of known sample vector pairs $\{(\mathbf{x}_k, y_k) |\ k=1,...,NUM\}$. In typical applications a suitable σ can be found quite easily by trial and error. This is because the relation between σ and MSE is usually smooth with a fairly broad minimal MSE section. Consequently, tuning is not overly critical to achieve an adequate overall model.

THE PREVIOUS TAPLR LOUDSPEAKER RESPONSE MODEL

The loudspeaker frequency response model is built up by first measuring the responses at a number of key frequencies in front of the loudspeaker. The TAPLR model is then used to smoothly interpolate the response between the measured responses over the whole frequency range, as was demonstrated in (Zaknich 2004). This model of the loudspeaker's magnitude response over frequency at a given point in space was developed by direct application of equation (1) with appropriate parameter definitions. The resulting equation (3) was developed to model the loudspeaker frequency response $|\hat{H}_j(\omega)|$ at a single spot in 3-D space j based on M_ω number of $1/m^{th}$ octave frequency response measurements.

$$|\hat{H}_j(\omega)| = \frac{\sum_{i=0}^{M_\omega} Z_i^{(\omega)} |H_i^{(j)}| f(\|\omega-\omega_i\|,\sigma_\omega)}{\sum_{i=0}^{M_\omega} Z_i^{(\omega)} f(\|\omega-\omega_i\|,\sigma_\omega)} \tag{3}$$

where:

$f(\|\omega-\omega_i\|,\sigma_\omega)$ is a suitable radial basis function, typically Gaussian.

$|\hat{H}_j(\omega)|$ is the speaker magnitude response at fixed location j in 3-D space.

ω is an arbitrary frequency.

ω_i is the i th $1/m^{th}$ octave measurement frequency, where $i = 0$ is reserved for 0 Hz at which $|H_0^{(j)}| \approx 0$ without need for measurement.

σ_ω is the smoothing parameter for the frequency variable.

$|H_i^{(j)}|$ is the response measurement made at ω_i at fixed a location j in 3-D space.

M_ω is the total No. of measurement frequencies ω_i.

$Z_i^{(\omega)}$ is the weight associated with each i for adequate frequency interpolation.

The $1/m^{th}$ octave response measurements $|H_i^{(j)}|$ are typically made at a fixed point j located at a standard distance of one metre from the loudspeaker. Once these measurements have been made the loudspeaker model $|\hat{H}_j(\omega)|$ is tuned by a suitable selection of variables σ_ω and $Z_i^{(\omega)}$, which are then fixed and remain suitable for all loudspeaker models at any other spatial location j. The weighting factors $Z_i^{(\omega)}$ are computed by equation (4), while σ_ω is selected by trial and error in such a way as to produce the minimum Mean Squared Error (MSE) between model outputs and a set of measured responses spread over the whole frequency range but not at frequencies ω_i. The M_ω measurement frequencies ω_i are computed by equation (5), given a nonzero first measurement frequency ω^0. Frequency $\omega_0 = 0$ ($Z_0^{(\omega)}=1$) need not be measured since the response at that frequency can be defined to be zero or arbitrarily small. Figure 2 shows a representative response curve and the nonlinear spacing between measurement frequencies. Each measurement frequency ω_i could either represent a single frequency or it could represent the centre frequency of a $1/m^{th}$ octave band of frequencies. $\left|H_i^{(j)}\right|$ in equation (3) is a constant and corresponds to the $l_i(\mathbf{x})$ in the TAPLR equation (1).

$$Z_i^{(\omega)} = \left[(2)^{\frac{1}{m}} \right]^{i-1}, \quad i = 1,..,M_\omega \tag{4}$$

$$\omega_i = \omega^0 \left[(2)^{\frac{1}{m}} \right]^{i-1}, \quad i = 1,..,M_\omega \tag{5}$$

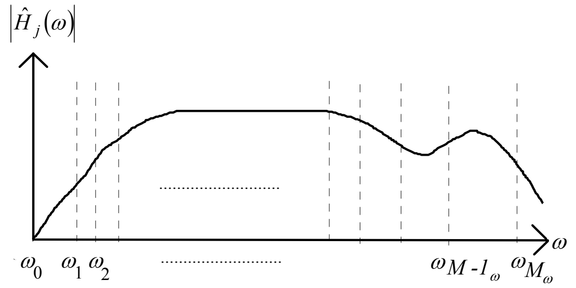

Figure 2. Representative TAPLR Speaker Model

Phase Response by Hilbert Transform

Loudspeakers are typically minimum phase systems and consequently their phase characteristics are directly related to their magnitude responses. The phase response of a minimum phase system $S(e^{j\omega T_s})$ can be reconstructed from the two-side magnitude response $\left|S(e^{j\omega T_s})\right|$ using the Hilbert Transform. This is commonly done in loudspeaker testing where it is relatively easy to measure the magnitude response but not the phase response, due to the uncertainty of the testing signal delay through the air. The phase response, $\angle S(e^{j\omega T_s})$, in radians is found by taking the real part of the Hilbert transform $\mathsf{H}\{\}$ of $\ln(\left|S(e^{j\omega T_s})\right|)$ as defined by equation (6).

$$\angle S(e^{j\omega T_s}) = real\{\mathsf{H}\{\ln(|\, S(e^{j\omega T_s})\,|)\}\} \tag{6}$$

If a Discrete Fourier Transform (DFT) is used to compute the Hilbert transform (Oppenheim and Schafer 1975) (Poularikas 1996) then the number of sample points N should be chosen such that NT_S is at least twice the length of the impulse response $s[nT_s]$. The phase in radians is found in the real part of the final result. Assume that a discrete Hilbert transform pair $u[n]$ and $v[n]$ is defined by $u[n] \overset{\mathsf{H}}{\longleftrightarrow} v[n]$, and a discrete Fourier transform pair $v[n]$ and $V(e^{j\omega T_s})$ is defined by $v[n] \overset{\text{DFT}}{\longleftrightarrow} V(e^{j\omega T_s})$.

Now, if $u[n] \overset{\text{DFT}}{\longleftrightarrow} U(e^{j\omega T_s})$ and $v[n] \overset{\text{DFT}}{\longleftrightarrow} V(e^{j\omega T_s})$,

then $v[n] \overset{\text{DFT}}{\longleftrightarrow} V(e^{j\omega T_s}) = -j\,\mathrm{sgn}(\omega) U(e^{j\omega T_s})$,

$$\text{where, } \mathrm{sgn}(\omega) = \begin{cases} +1 & 0 < \omega < \frac{\omega_2}{2} \\ 0 & \omega = \frac{\omega_2}{2} \\ -1 & -\frac{\omega_2}{2} < \omega < 0 \end{cases}$$

It therefore follows that,

$$u[n] \overset{\text{DFT}}{\longleftrightarrow} U(e^{j\omega T_s}) \rightarrow V(e^{j\omega T_s}) = -j\,\mathrm{sgn}(\omega) U(e^{j\omega T_s}) \overset{\text{DFT}^{-1}}{\rightarrow} v[n].$$

For this application equation (3) is taken and uniformly frequency sampled to provide the required two-side magnitude response $\left|S(e^{j\omega T_s})\right|$ from which the associated phase response is computed.

REVIEW OF OLD TAPLR MODEL TEST RESULTS AND EXPERIMENTAL SETUP

The original TAPLR model was tested using a small 11.6 litre loudspeaker box system having a woofer and tweeter combination plus associated frequency compensation circuitry (Zaknich 2004). This box was placed centrally 1 m in front of a calibrated microphone in a room with a floor to ceiling height of 2.8 metres, as shown in Figure 3.

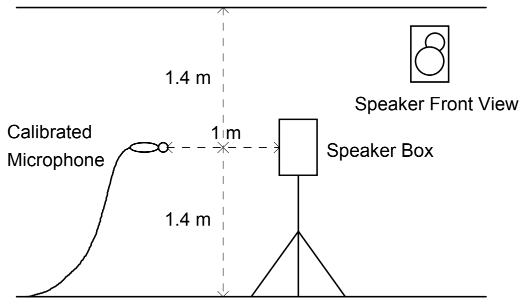

Figure 3. Experimental Setup

The test signal for each $1/m^{th}$ octave frequency band centred at each ω_i was a short linear frequency swept burst whose frequency ranged between the neighbouring $1/m^{th}$ octave frequency band edges. This signal was made to be either at least one full cycle or just over 5 ms in length. For frequencies below 200 Hz it was one cycle beginning and ending at zero amplitude and for frequencies above 200 Hz it was a complete number of cycles, being at least 5 ms and completing its last cycle after that. By using this signal it was possible to accurately measure the direct frequency response, avoiding any room reverberation errors, for frequencies above about 300 Hz and minimise the errors for frequencies lower. Good models can also be made by using similarly designed constant frequency bursts at the $1/m^{th}$ octave frequency values.

The response for each band was calculated by taking the ratio of the RMS values of the transmitted and received signals using a digital signal processing system running at a sampling rate of 48 KHz and using 16 bit digital to analogue and analogue to digital converters. The loudspeaker's response was also measured by taking an impulse response measurement, but this was only accurate for frequencies above about 300 Hz due to room reverberation effects. The burst measurements were taken at standard $^1/_{12th}$ octave bands beginning with the first band frequency of $\omega^0 = 2\pi\ 15.84893192$ radians per second and using M_ω = 128 bands to $\omega_{M_\omega} = 2\pi\ 23714$ radians per second.

The frequency variable ω in equations (3) and (5) is normally defined in radians per second but with no loss of generality all the following tests were done taking it to be in Hz. Figure 4 compares the frequency response measured using the $^1/_{12th}$ octave linear bursts with the standard impulse response. The $^1/_{12th}$ octave measurements were somewhat corrupted by room reverberation effects below about 300 Hz but it is accurate enough for present purposes. For example, the sharp peak at 200Hz is an artefact of ceiling and floor bounce. The impulse response was shortened to about 5 ms length to remove the reverberation effects and is therefore only accurate above 300Hz. It can be seen that the two methods track very closely for frequencies above about 300Hz.

Figure 4. Burst and Impulse Responses Measurements

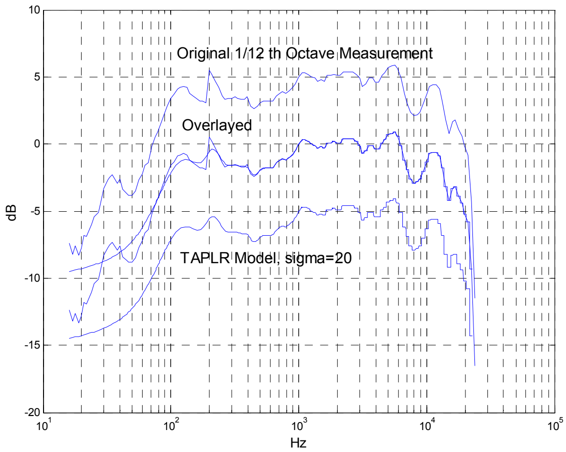

Figure 5. TAPLR Model Magnitude, Log Scale

The mean squared error difference between the normalised $^1/_{12th}$ octave measurements and the normalised impulse response was 0.0053 over the frequency range 400 to 16,000 Hz, which translates to +/- 0.61 dB error. This is well within the accepted experimental error of +/- 1 dB for acoustic measurements and demonstrates that the $^1/_{12th}$ octave measurements are acceptably accurate for all practical purposes.

When the TAPLR model is used to cover the whole frequency range using a σ_ω = 20 it can be seen in Figure 5 that the difference between this model and the $^1/_{12th}$ octave measurements is very small (much smaller than +/- 0.61 dB). However, the high frequency end of the model does tend toward a stair step response. Although the error is negligible for practical purposes it would still be desirable to have a perfectly smooth model over the whole frequency band to avoid aliasing errors in digital realisations due to sharp transitions in the frequency response. This was achieved in (Zaknich 2004) by breaking the frequency into 4 arbitrary bands (LF = $0 \rightarrow 999$Hz, MF1 = $1000 \rightarrow 4999$Hz, MF2 = $5000 \rightarrow 9999$Hz, HF = $10000 \rightarrow 24000$Hz) and applying four separate TAPLR models with different σ_ω values of 10, 100, 200 and 500 respectively.

Smaller models with less resolution and accuracy, but still acceptable for some applications can be developed using $^1/_{6th}$ or $^1/_{3th}$ octave measurements, i.e., every second or fourth value of the $^1/_{12th}$ octave measurements respectively.

THE NEW MODEL

Using TAPLR equations (1) and (3) as a basis for the loudspeaker response model is only appropriate when the ratio of smallest to the largest frequency difference between measured neighbouring frequency centres ω_i is less than about 2 to 10, as has been approximately achieved in (Zaknich 2004) by breaking the range into the 4 bands. This is because the Gaussian RBF reduces very quickly to zero away from its centre, relative to the chosen σ_ω value, and there is a loss of effective interpolation between centres too far away relative to the σ_ω value. What happens in that case is that the nearest frequency centre, and its associated measured response value $\left|H_i^{(j)}\right|$, to the test frequency ω is chosen as the output, producing the stair step effect. This problem suggested an obvious improvement from that of the model equation (3) to that of equation (7), where the weighting $Z_i^{(\omega)}$ is now applied to the RBF's sigma σ_ω rather than to the RBF functional value itself. In this way the width of each RBF is made to be relatively proportional to the distances between neighbouring RBF centres ω_i .

$$\left|\hat{H}_j(\omega)\right| = \frac{\sum_{i=0}^{M_\omega} \left|H_i^{(j)}\right| f(\left\|\omega - \omega_i\right\|, Z_i^{(\omega)}\sigma_\omega)}{\sum_{i=0}^{M_\omega} f(\left\|\omega - \omega_i\right\|, Z_i^{(\omega)}\sigma_\omega)} \tag{7}$$

It is this new model that is tested in this paper for accuracy and effectiveness and compared with the previous TAPLR model. Figures 6 and 7 show the comparisons between the original one-twelfth octave magnitude measurement and the modified TAPLR magnitude model on logarithmic and linear scales respectively. Different σ_ω selections of 0.1, 0.7 and 2.0 in Figure 6 show how the variation in σ_ω affects model accuracy. For small σ_ω values the model is a stair step model but unlike the old TAPLR model equation (3) a very close and faithful representation is seen from low to high frequencies. As σ_ω is increased the model simply gets smoother but it still holds faithfully to the underlying measured curve over the whole frequency range. From this it can be seen that a suitable selection of $\sigma_\omega = 0.7$ is not overly critical to achieving an accurate model that is certainly well within a maximum error +/- 1 dB against both the original one-twelfth octave magnitude measurement and impulse responses (see Figure 4).

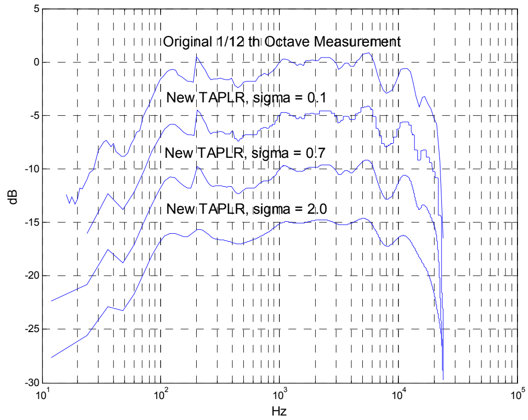

Figure 6. New TAPLR Model Magnitude, Log Scale

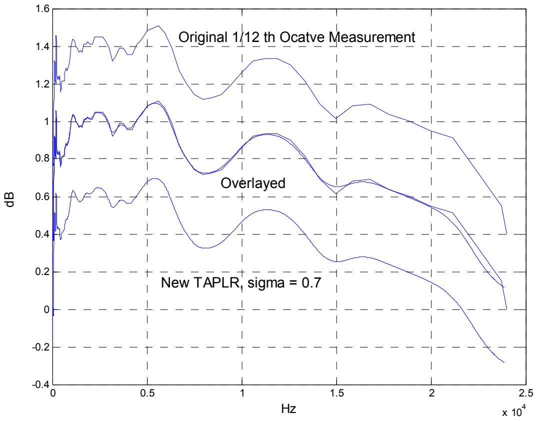

Figure 7. New TAPLR Model Magnitude, Linear Scale

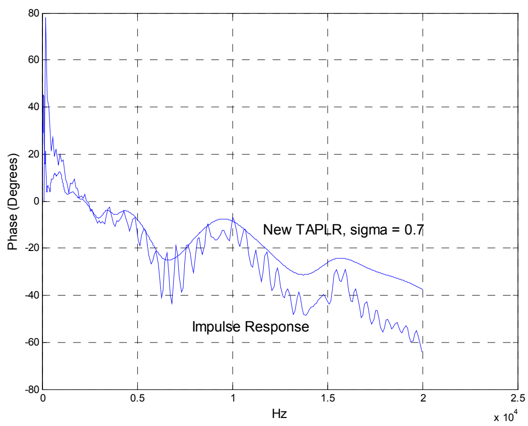

Figure 8. Phase Angle Responses

The loudspeakers phase response was computed by the Hilbert transform method described above using the new modified TAPLR magnitude model with $\sigma_\omega = 0.7$. This phase response can be seen in Figure 8, where it is compared with the phase response of the impulse response measurement, using the same Hilbert transform method. They differ slightly because the impulse response based model has a lower low frequency magnitude response due to the truncation of the impulse response sequence, done to avoid room reverberations. Nevertheless, they do show good general correspondence at the higher frequencies with the new TAPLR model having a smaller phase angle dip because it has a broader bandwidth.

CONCLUSIONS

The advantage of this new TAPLR loudspeaker frequency response model (7) is that it can provide a smooth and continuous (differentiable) interpolation model within a single equation, which is loaded instantly from a finite number of direct local response measurements. This is possible because the width of each RBF is now made to be proportional to the distances between neighbouring RBF centres ω_i, providing accurate interpolation weightings along the whole frequency range. The equation (7) can be made a little more computationally efficient by involving only a few of the neighbouring RBF measurements $\left|H_i^{(j)}\right|$ around the desired frequency ω. This is possible because the tails from far away RBFs will have negligible contribution toward the computation, but it is hardly necessary because the computational burden is not so great. There are other interpolation methods that could be used, including polynomial fitting, quadratic spline fitting and piecewise linear interpolation but these are either more complex to do or they do not provide as smooth a result.

The new model's advantages become more apparent if the loudspeaker is to be statically or dynamically equalised. The equalisation can be achieved easily by just computing the corrections at the finite number of $^1/_{m}th$ octave measurement frequencies and leaving the model to interpolate all the required frequency responses in between. This is much more convenient to do than to solve the interpolation curve a new every time a change in equalisation is required.

REFERENCES

J. Moody, and C. Darken, 1989, *Fast learning in networks of locally-tuned processing units*, Neural Computation, Vol. 1 No. 2, pp. 281-294.

A.V. Oppenheim, and R.W. Schafer, 1975, *Digital signal processing*, Prentice-Hall.

A. Poularikas, Editor in Chief, 1996, *The transforms and applications handbook*, CRC and IEEE Press.

D. F. Specht, 1991, *A general regression neural network*, IEEE Transactions on Neural Networks, Vol. 2, No. 6, pp. 568-576.

A. Zaknich, 1998, *Introduction to the modified probabilistic neural network for general signal processing applications*, IEEE Transactions on Signal Processing, Vol. 46, No. 7, pp. 1980-1990.

A. Zaknich, and Y. Attikiouzel, 2000, *A tuneable approximate piecewise linear model derived from the modified probabilistic neural network*, IEEE Signal Processing Workshop on Neural Networks for Signal Processing (NNSP), Sydney, Australia, Vol. 1, pp. 45-53.

A. Zaknich, 2004, *Loudspeaker response model using tuneable approximate piecewise linear regression*, IEEE International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, Vol. 4, pp. 2711-2716.